Trust Prediction with Propagation and Similarity Regularization

Xiaoming Zheng[1], Yan Wang[1], Mehmet A. Orgun[1], Youliang Zhong[1], Guanfeng Liu[2]
[1]Department of Computing, Macquarie University, Sydney, NSW 2109, Australia
{xiaoming.zheng, yan.wang, mehmet.orgun, youliang.zhong}@mq.edu.au
[2]School of Computer Science and Technology, Soochow University, Suzhou, Jiangsu, China 215006
gfliu@suda.edu.cn

Abstract

Online social networks have been used for a variety of rich activities in recent years, such as investigating potential employees and seeking recommendations of high quality services and service providers. In such activities, trust is one of the most critical factors for the decision-making of users. In the literature, the state-of-the-art trust prediction approaches focus on either dispositional trust tendency and propagated trust of the pair-wise trust relationships along a path or the similarity of trust rating values. However, there are other influential factors that should be taken into account, such as the similarity of the trust rating distributions. In addition, tendency, propagated trust and similarity are of different types, as either personal properties or interpersonal properties. But the difference has been neglected in existing models. Therefore, in trust prediction, it is necessary to take all the above factors into consideration in modeling, and process them separately and differently.

In this paper we propose a new trust prediction model based on trust decomposition and matrix factorization, considering all the above influential factors and differentiating both personal and interpersonal properties. In this model, we first decompose trust into trust tendency and tendency-reduced trust. Then, based on tendency-reduced trust ratings, matrix factorization with a regularization term is leveraged to predict the tendency-reduced values of missing trust ratings, incorporating both propagated trust and the similarity of users' rating habits. In the end, the missing trust ratings are composed with predicted tendency-reduced values and trust tendency values. Experiments conducted on a real-world dataset illustrate significant improvement delivered by our approach in trust prediction accuracy over the state-of-the-art approaches.

1 Introduction

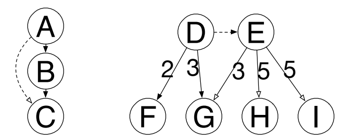
In recent years, a diverse range of online social networks (OSNs), such as Facebook, MySpace, Twitter, LinkedIn and Google+, have attracted an increasingly large number of users. Moreover, OSNs have proliferated to be the platforms for a variety of rich activities, such as investigating potential employees as well as seeking recommendations of high quality services and service providers. For example, according to a survey on 2600 hiring managers in 2009 by CareerBuilder[1], 45% of those managers used social networking sites to investigate potential employees, and in 2013, the ratio increased to 92%. In the context of such activities, *trust*, the commitment to a future action based on a belief that it will lead to a good outcome (Golbeck and Hendler 2006), is one of the most critical factors for the decision making of users. This demands effective approaches and mechanisms to predict the trust between two users without any direct connection.

Trust prediction is the process of estimating a new pair-wise trust relationship between two users who are not directly connected based on existing observations. In the literature, there are basically two groups of trust prediction approaches: propagation based trust prediction (i.e., trust propagation/inference) and similarity based trust prediction. *Trust propagation/inference* is the process of evaluating trust from a source user to a target user along a path between them that consists of links and trust values (Guha et al. 2004). For example, as shown in Fig. 1(a), if user A trusts user B, and user B trusts user C, then A trusts C to some extent (Golbeck and Hendler 2006; Liu, Wang, and Orgun 2009). Trust propagation has been studied in many web application areas including e-commerce (Wang and Lin 2008; Zhang, Wang, and Zhang 2012b; 2012a), Peer-to-Peer systems (Xiong and Liu 2004), and social networks (Jøsang and Ismail 2002; Golbeck and Hendler 2006; Liu et al. 2013). On the other hand, a user tends to trust other users who are similar to himself/herself (Lin et al. 2012). Broadly speaking, *similarity based trust prediction* is to estimate the trust between two users from their similar habits, context and profiles.

In the literature, similarity is calculated from two users' common trust rating values given to others (Ma et al. 2011a). Such similarity is termed as *trust rating value similarity* in this paper. In the meantime, it should be noted that similarity can also be calculated from two users' distributions of trust ratings, which is termed as *trust rating distribution similarity*. The distribution of trust ratings offers a new way of identifying users' different behaviors and improving trust prediction accuracy. For example, as shown in Fig. 1(b), the trust values given to G by D and E are the same. However, they

Copyright © 2014, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

[1]http://www.careerbuilder.com



(a) Trust propagation (b) Trust Ratings

Figure 1: Examples

come from two different distributions showing that even the same trust value given to G could be different in the minds of D and E — the trust value 3 is the higher value user D has given while it is the lower one in E's ratings.

Propagated trust and the two types of similarities are interpersonal properties. They are the factors that influence the trust between two users. By contrast, *trust tendency* (also termed as *trust bias* in (Yao et al. 2012; 2013b)) is a type of factors extracted from all the trust ratings that one user gave or received, showing his/her dispositional tendency to trust others or to be trusted by others on average (termed as *truster tendency* and *trustee tendency* respectively) (Yao et al. 2012; 2013b). Trust tendency is regarded as a very important concept in Social Science and it is recognized as an integral part of the final trust decision (Tversky and Kahneman 1974). For instance, some users tend to give relatively high trust ratings more generously than others while some users receive higher trust ratings compared with others. The details of trust tendency will be presented in Section 4.

In the literature, the existing works predict trust either via trust propagation only (Guha et al. 2004; Golbeck and Hendler 2006; Liu, Wang, and Orgun 2009; Wang, Li, and Liu 2013), or considering propagated trust and tendency (Yao et al. 2013a; 2013b), or merely utilizing the similarity of rating values (Ma et al. 2011a; 2011b). In general, they have the following drawbacks. *First*, all the tendency, propagated trust and similarity influence the trust between two users. All of them should be utilized to predict pair-wise trust, rather than considering one or two influential factors only. *Second*, the similarity of trust rating distributions describes the similarity of users' behaviors in giving trust ratings. Thus, it is valuable for trust prediction (Zheng, Wang, and Orgun 2013). However, it has been neglected in the literature. *Third*, all these factors are of different types representing either personal properties or interpersonal properties. Therefore, they should be processed separately and differently so as to deliver high accuracy in trust prediction.

In order to overcome the above drawbacks, in this paper, we propose a new trust prediction model based on rating decomposition and matrix factorization, incorporating both propagation based trust prediction and similarity based trust prediction. The main contributions of our work are summarized as follows:

1. Our model decomposes trust ratings into trust tendencies (i.e., truster tendency and trustee tendency) and tendency-reduced ratings, and predicts trust with tendency-reduced ratings to reduce the negative effect of trust tendency.
2. Our model considers the similarity of trust rating distributions to further differentiate the trust between users and optimize matrix factorization. This is particularly important when the similarity of trust rating values is the same.
3. Our model considers both propagated trust and similarity factors, which consist the propagation and similarity regularization term of matrix factorization, in order to improve the trust prediction accuracy.
4. Based on the commonly used metrics of Root Mean Square Error (*RMSE*) and Mean Absolute Error (*MAE*), the experiments conducted on a real-world dataset have demonstrated significant improvements delivered by our model in trust prediction accuracy over the state-of-the-art approaches.

2 Related Work

In this section, we briefly review the related works in three areas, including trust propagation, collaborative filtering and matrix factorization.

2.1 Trust Propagation

In the past decade, most of the trust prediction models have focused on the propagation of trust along the paths connecting users. Guha et al. (Guha et al. 2004) propose a trust propagation model considering the number of hops when calculating the propagated trust value between a source user and the target one. Golbeck et al. (Golbeck and Hendler 2006) propose a trust propagation approach to establish a trust relationship between two indirectly connected users based on the average trust value along a social path from the source user to the target user. Huang et al. (Hang, Wang, and Singh 2009) utilize operators such as concatenation, aggregation and selection to propagate trust. Liu et al. (Liu, Wang, and Orgun 2010) argue that social relationships and recommendation roles are also important for trust propagation. These approaches have improved the accuracy of trust prediction considerably. But many other factors influencing trust prediction should be taken into account, such as tendency and similarity.

2.2 Collaborative Filtering

In trust prediction, if trustees (the users who receive trust ratings) are regarded as the items in a recommender system, the Collaborative Filtering (CF) approach employed in recommender systems can be leveraged (Yao et al. 2013a; 2013b). As such CF approaches analyze the relationships between trusters (the users who give trust ratings) and trustees to identify new truster-trustee associations (Koren, Bell, and Volinsky 2009). They mainly rely on users' past behaviors, like previous ratings. Typically, there are two types of CF: neighborhood methods and latent factor methods. The former predicts missing ratings based on ratings of similar neighbors. The latter explains the ratings by featuring both trusters and trustees on a number of latent factors inferred from ratings. Most successful realizations of latent factor models are based on the foundation of matrix factorization. In Netflix price competition, Koren et al. (Koren, Bell, and Volinsky 2009) have shown that matrix factorization methods outperform other rating prediction methods significantly, especially in sparse datasets. Salakhutdinov et al. (Salakhutdinov and Mnih 2008b) have shown that the matrix factorization based method could scale well to large datasets as it scales linearly with the number of observations and performs well on very sparse and imbalanced datasets.

2.3 Modifications to Matrix Factorization

Matrix factorization methods have also been modified in different ways to improve prediction accuracy. Ma et al. (Ma et al. 2011a) incorporate social regularization into matrix factorization achieving better accuracy. Yao et al. (Yao et al. 2013b) modify matrix factorization by regarding tendency and propagated trust as some latent factors of the basic matrix factorization to boost trust prediction accuracy.

However, with tendency and propagated trust values only, trust can not be predicted accurately when there is no existing path. Zheng et al. (Zheng, Wang, and Orgun 2013) have shown that the distribution of users' trust ratings is an important factor that influences the trust between the source user and the target user. Therefore, it is essential to take advantage of distribution to boost trust prediction accuracy further.

In addition, the way in which the influential factors are used needs to be improved as well. On one hand, personal factors such as tendencies are decomposed from every single user's ratings and influence the user's global ratings; on the other hand, interpersonal factors such as similarity and propagated trust are extracted from two users' trust ratings to reflect the features between them. Therefore, the two types of factors should be treated differently in order to improve trust prediction accuracy.

Different from the existing approaches, our approach decomposes trust ratings into truster tendency, trustee tendency and tendency-reduced ratings. Based on tendency-reduced ratings we extend matrix factorization methods by adding a propagation and similarity regularization term which incorporates propagated trust, rating value similarity and distribution similarity to put constraints on the difference between two users' latent feature vectors. In particular, an important feature of our approach is that we do not impose any limitation on latent factors of matrix factorization. As we will show later in the paper, our approach boosts the prediction accuracy of trust ratings significantly.

3 Basic Matrix Factorization

In this section, we present the basic matrix factorization method from the viewpoint of trust prediction. Matrix factorization is an efficient and effective approach in recommender systems to factorizing the user-item rating matrix into user-specific and item-specific matrices and predicting missing data based on both matrices (Ma et al. 2011a; Salakhutdinov and Mnih 2008a; 2008b). In the application of trust prediction, trustees are regarded as the "items" in recommender systems (Yao et al. 2013a). Thus, matrix factorization methods factorize the trust ratings matrix into truster-specific and trustee-specific matrices respectively.

We consider an $n \times n$ trust rating matrix R describing n trusters' numerical ratings on n trustees. The matrix factorization models map both trusters and trustees to a joint latent factor space of dimensionality l, so that truster-trustee trust ratings are modeled as inner products in that space. Accordingly, each truster i is associated with a vector $u_i \in \mathbb{R}^l$, while each trustee is associated with a vector $v_j \in \mathbb{R}^l$. Finally, all the vectors $\{u_i\}$ constitute the truster-specific matrix U indicating to what extent the corresponding users trust others w.r.t. the specific latent factors. Meanwhile, vectors $\{v_j\}$ compose the trustee-specific matrix V indicating to what extent the corresponding users are trusted by others w.r.t. the specific latent factors. So, the rating matrix R is factorized as a multiplication of l-rank factors,

$$R \approx U^T V, \tag{1}$$

where $U \in \mathbb{R}^{l\times n}$ and $V \in \mathbb{R}^{l\times n}$ with $l < n$. Once the factorization is completed, the missing ratings could be calculated from

$$r_{i,j} \approx u_i^T v_i. \tag{2}$$

Note that user u_i and user v_i are the same user with two different roles—truster and trustee respectively. The factorization is achieved by minimizing the equation:

$$\frac{1}{2}||R - U^T V||_F^2, \tag{3}$$

where $||.||_F^2$ represents the Frobenius norm. Note that each user only gives trust ratings to a few other users. Hence, the matrix R contains a large amount of missing values as an extremely sparse matrix. Therefore, Eq. (3) is changed to

$$\min_{U,V} \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} I_{ij}(r_{ij} - u_i^T v_j)^2, \tag{4}$$

where I_{ij} is an indicator function. $I_{ij} = 1$ *iff.* user i (truster) rated user j (trustee) , $i \neq j$. Otherwise, $I_{ij} = 0$. In order to avoid overfitting, two regularization terms from zero-mean spherical Gaussian priors (Salakhutdinov and Mnih 2008b) are placed into Eq. (2). Hence, we have

$$\min_{U,V} \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} I_{ij}(r_{ij} - u_i^T v_j)^2 + \frac{\lambda_1}{2}||U||^2 + \frac{\lambda_2}{2}||V||^2, \tag{5}$$

where $\lambda_1 > 0$ and $\lambda_2 > 0$. Thus, the learning process of the method can be achieved by Eq. (5) using the gradient descent method (Koren, Bell, and Volinsky 2009).

4 Our Proposed Trust Prediction Approach

In this section, we first discuss the factors influencing trust between users in detail. Then we propose a novel method to incorporate these influential factors into a regularization in matrix factorization to improve trust prediction accuracy.

4.1 Factors that influence trust

In real life, trust is influenced by many factors, including trust tendency, propagated trust, value similarity and distribution similarity (Golbeck and Hendler 2006; Ma et al. 2011a; Liu, Wang, and Orgun 2012; Jia, Zhang, and Liu 2013; Yao et al. 2013b; Zhang, Wang, and Zhang 2013).

Trust Tendency: When rating others in trust, some users give relatively higher trust ratings than the others, showing different tendencies. On the other hand, some users receive higher trust ratings than others, meaning that they are more likely to be trusted. So, there are two types of tendencies in trust ratings: *truster tendency* $T_u(i)$ and *trustee tendency* $T_v(i)$. Truster tendency can be considered as a personal property that implies a user's average dispositional tendency to trust others. It can be calculated as the average of all the trust ratings a user i gives to others (Yao et al. 2013a). On the other hand, trustee tendency can be treated as another personal property that shows a user's tendency to be trusted. It can be calculated as the average of all the ratings a user j received (Yao et al. 2013a). $\hat{r}_{ij}$ is the trust ratings adjusted by both tendencies above termed as *tendency-reduced ratings* in the paper. Therefore, each trust rating r_{ij} can be decomposed as $r_{ij} = \alpha_1 T_u(i) + \alpha_2 T_v(j) + \alpha_3 \hat{r}_{ij}$, where, α's are the coefficients. Only $\hat{r}_{ij}$ is used for matrix factorization. Thus, the negative effect of trust tendency can be reduced.

Propagated Trust: It is concluded in social network studies that people can trust a stranger to some extent if the person is a friend's friend (Guha et al. 2004). Thus, many trust propagation methods infer trust along a path between two users without direct connections (Golbeck and Hendler 2006; Liu et al. 2013). Here, we adopt the propagation method introduced in (Liu et al. 2013; Li, Wang, and Lim 2009) to select the path with the highest propagated trust value $infer(i, j)$ between user i and user j by multiplication within H hops. If no path is available within H hops, we set $infer(i, j) = 0$. Here, $infer(i, j) \neq infer(j, i)$ in most circumstances because when user i trusts user j with a certain trust value, it does not mean user j trusts user i to the same extent.

Trust Rating Value Similarity: Conventionally, with the rating information of all the users, the trust rating value similarity of two users can be calculated from the common trust ratings that the two users give to others (Jia, Zhang, and Liu 2013). The most prevalent approaches of this similarity evaluation are Vector Space Similarity (VSS) and Pearson Correlation Coefficient (PCC) (Breese, Heckerman, and Kadie 1998). VSS calculates the similarity from ratings of common trustees that user i and user j have rated respectively:

$$vss(i,j) = \frac{\sum\limits_{f \in I(i) \bigcap I(j)} r_{if} \cdot r_{jf}}{\sqrt{\sum\limits_{f \in I(i) \bigcap I(j)} r_{if}^2} \cdot \sqrt{\sum\limits_{f \in I(i) \bigcap I(j)} r_{jf}^2}}, \quad (6)$$

where user f belongs to the subset of trustees that user i and user j both have rated. r_{if} and r_{jf} are the trust ratings user i and user j give to user (trustee) f.

On the other hand, PCC takes into account the rating styles that some users would like give relatively higher ratings to all the others while some may not. Hence, PCC adds a mean of ratings as follows:

$$pcc(i,j) = \frac{\sum\limits_{f \in I(i) \bigcap I(j)} (r_{if} - \overline{r}_i) \cdot (r_{jf} - \overline{r}_j)}{\sqrt{\sum\limits_{f \in I(i) \bigcap I(j)} (r_{if} - \overline{r}_i)^2} \cdot \sqrt{\sum\limits_{f \in I(i) \bigcap I(j)} (r_{jf} - \overline{r}_j)^2}}, \quad (7)$$

where $\overline{r}_i$ and $\overline{r}_j$ represent the average rates of user i and user j respectively. In addition, the range of the PCC is $[-1, 1]$. Thus, PCC is normalized into $[0, 1]$ in applications by $\mathfrak{q}(x) = (\mathfrak{p}(x) + 1)/2$ (Ma et al. 2011a).

Trust Rating Distribution Similarity: The distribution of a user's ratings reveals the user's rating habits. For example, a user gives diverse ratings with equal probability (Uniform distribution) while another user prefers giving a certain trust rating value with a high probability (Gaussian distribution). The same trust value from these two distributions should be treated differently. Kullback-Leibler (KL) -distance (Relative Entropy) is a natural distance function from one user's distribution of ratings to the other's (Koller and Friedman 2009). It can depict the difference in trust rating distributions between two users. For discrete probability distributions, the KL-distance is formulated as follows:

$$D_{KL}(i||j) = \sum_k \ln\left(\frac{P_i(k)}{P_j(k)}\right) P_i(k), \quad (8)$$

where $k \in K$ is the space of all the trust ratings that user i has given; P_i and P_j are the trust rating distributions of users i and j. As the range of KL-distance is $[0, \infty]$, we use the projection function $q(x) = e^{-p(x)}$ to convert the range to $[0, 1]$, where, after conversion, 1 means the two distributions are exactly the same while 0 means they are different.

Different from trust tendency, the last three factors have the same characteristics that they influence the trust between two users and have the same value range and trend (after conversion). The weighted sum of interpersonal trust factors between user i and user j is termed as *trust factor utility*, which can be formulated as:

$$TF(i,j) = \beta_1 infer(i,j) + \beta_2 pcc(i,j) + \beta_3 D_{KL}(i||j) \quad (9)$$

where β's are the coefficients.

4.2 Our Modified Matrix Factorization

As mentioned above, studies in Social Science have pointed out that people would like to seek suggestions from friends in the real world. They adopt suggestions according to the trust levels of friends which are influenced by interpersonal trust factors (Berscheid and Reis 1998). Hence, we propose a propagation and similarity regularization term to impose constraints between truster i and trustee f to minimize the distances between user-specific vectors u_i and u_f. It is formulated as:

$$\frac{\gamma}{2} \sum_{i=1}^{n} \sum_{f \in \mathcal{F}^+(i)} TF(i,f) ||u_i - u_f||_F^2, \quad (10)$$

where $\gamma > 0$, $\mathcal{F}^+(i)$ is the set of trustees who, at least, have a trust path connected from truster i. $TF(i, f)$ is the trust factor utility in Eq. (9). If a trustee $f \in \mathcal{F}^+(i)$ of user i has a very similar habit to i and a high trust value propagated from user i, then the value of $TF(i, f)$ will be close to 1, otherwise it is close to 0. Furthermore, a small value of $TF(i, f)$ means that the distance between user-specific vectors u_i to u_f should be large while a large value of $TF(i, f)$ indicates the distance should be small. Thus, the trust factor utility $TF(i, j)$ enables the matrix factorization method to incorporate the different similarities and propagated trust between user i and his/her truster or trustee. Finally, our trust prediction model can be formulated as:

$$\begin{aligned}\min_{U,V} \mathcal{L}(R,U,V) = &\min_{U,V} \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} I_{ij}(\hat{r}_{ij} - u_i^T v_j)^2 \\ &+ \frac{\lambda_1}{2}||U||^2 + \frac{\lambda_2}{2}||V||^2 \\ &+ \frac{\gamma}{2}\sum_{i=1}^{n}\sum_{f\in\mathcal{F}^+(i)} TF(i,f)||u_i - u_f||_F^2.\end{aligned} \tag{11}$$

In our model, $TF(i,j) \neq TF(j,i)$ in most cases, because trust propagation and KL-distance are asymmetric ($infer(i,j) \neq infer(j,i)$ and $D_{KL}(i||j) \neq D_{KL}(j||i)$) in most circumstances, indicating that "user i trusts user j" does not mean "user j trusts user i to the same extent".

This method improves the accuracy of trust prediction and propagates users' trust ratings indirectly. In details, if user i rates user f and user f rates user g (suppose user i does not rate user g), the distances between feature vectors u_i and u_g is minimized when we minimize $TF(i,f)||u_i - u_f||_F^2$ and $TF(f,g)||u_f - u_g||_F^2$.

A local minimum value of the objective function (11) can be obtained using gradient descent methods in latent factors of u_i and v_i:

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial u_i} = &-\sum_{j=1}^{n} I_{ij}(\hat{r}_{ij} - u_i^T v_j)v_j + \lambda_1 u_i \\ &+ \gamma\sum_{f\in\mathcal{F}^+(i)} TF(i,f)(u_i - u_f) \\ &+ \gamma\sum_{g\in\mathcal{F}^-(i)} TF(g,i)(u_i - u_g),\end{aligned} \tag{12}$$

$$\frac{\partial\mathcal{L}}{\partial v_i} = -\sum_{i=1}^{n} I_{ij}(\hat{r}_{ij} - u_i^T v_j)u_j + \lambda_2 v_i. \tag{13}$$

where $\mathcal{F}^-(i)$ is the set of trusters, each of whom, at least, has a trust path to trustee i.

5 Experiments

In this section, we present and analyze the results of the experiments we have conducted on a real-world dataset to illustrate the trust prediction accuracy of our method in comparison with the state-of-the-art approaches.

5.1 Dataset description

The dataset Advotago[2] used in our experiments is obtained from a trust-based social network. The network collects trust data between users and refreshes the dataset regularly. We adopt the dataset released on September 10th, 2013. It contains 7,425 users, among which 4,107 are trusters giving trust ratings to 4,699 trustees. In the dataset, there are 56,535 trust ratings given by 6,633 trusters, out of which 51,400 are pair-wise ratings and 5,135 are self-ratings. This paper aims to predict pair-wise ratings, and thus self-ratings are ignored. Trust ratings in this dataset are divided into 4 levels which are 'Observer', 'Apprentice', 'Journeyer' and 'Master' in ascending order. 'Observer' is the lowest trust level while 'Master' is the highest trust level. In our experiments, we map the trust levels of 'Observer', 'Apprentice', 'Journeyer' and 'Master' to 0.1, 0.4, 0.7 and 1 respectively.

5.2 Measures

In the area of prediction and recommendation, both the Mean Absolute Error (*MAE*) and the Root Mean Square Error (*RMSE*) are the most common metrics used to measure the prediction accuracy of a model (Yao et al. 2012; 2013b). Thus, we adopt them to compare the prediction accuracy of our proposed approach with the related state-of-the-art approaches. The metric *MAE* is formulated as:

$$MAE = \frac{1}{T}\sum_{i,j} |r_{ij} - \tilde{r}_{ij}|, \tag{14}$$

where r_{ij} denotes the actual trust ratings user i gives to user j. $\tilde{r}_{ij}$ represents the predicted trust ratings that user i will give to user j. T denotes the total number of trust ratings in the validation dataset. *MAE* weights the individual differences equally as a linear score.

The metric *RMSE* is defined as:

$$RMSE = \sqrt{\frac{1}{T}\sum_{i,j}(r_{ij} - \tilde{r}_{ij})^2}. \tag{15}$$

RMSE gives higher weights to larger errors as the errors are squared before taking their average. It is always larger or equal to the *MAE*. Both *MAE* and *RMSE* are usually used together to diagnose the variation in the errors of prediction. Lower values of *MAE* and *RMSE* mean better accuracy.

5.3 Comparisons

In order to evaluate the prediction accuracy of our approach, we compare it with two recent and promising approaches–an trust inference model (MATRI) (Yao et al. 2013a) and matrix factorization with social regularization (MFISR) (Ma et al. 2011b).

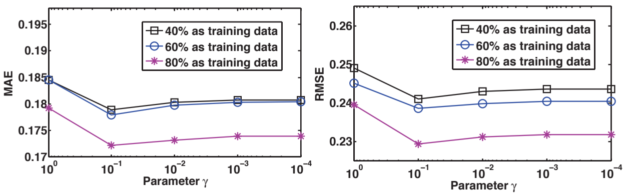
MATRI: This approach (Yao et al. 2013a; 2013b) considers trust tendency and propagated trust to predict missing trust ratings. The factors calculated from trust tendency and propagated trust are treated as some of the latent factors when conducting matrix factorization while other latent factors in matrix factorization are kept unchanged.

MFISR: This approach (Ma et al. 2011b) adds social regularization into conventional matrix factorization by introducing average-based and individual-based social regularization terms separately. In addition, matrix factorization with individual-based social regularization (MFISR) was proved to be more effective and accurate than that with average-based regularization. Therefore, in our experiments, we compare our method with MFISR.

5.4 Experimental Settings

In our model, the coefficients α's and β's determine the weight of each factor that influences the trust between two users. They are essential to the trust prediction accuracy. In

[2]http://www.trustlet.org/wiki/advogato_dataset.



(a) MAE (b) RMSE

Figure 2: Impact of Parameter γ

order to obtain the best coefficients, we treat each coefficient as a 'gene' and construct a 'chromosome' containing all the 6 coefficients of α's and β's. The prediction result from our modified matrix factorization is used as a fitness function. Thus, the real-valued Genetic Algorithm (Mitchell 1998) has been used to determine the best weight for each trust factor. To make comparison fair, this method is also used for both MATRI and MFISR to determine their coefficients on the same training dataset.

In total, we have conducted three groups of experiments with different percentages (80%, 60% and 40%) of the data for training. 10 groups of randomly generated initial matrices are used to initialize each model. In all of the three approaches, we use the same gradient descent method for the matrix factorization process and set $\lambda_1 = \lambda_2 = 0.01, \gamma = 0.1$, $H = 2$ and $l = 10$.

Setting of parameters γ and λ: γ and λ's are very important factors. γ controls to what extent propagation and similarity regularization should be incorporated while λ's manage to what extent Gaussian priors should be incorporated. As the impact of λ's share the same trend as γ in terms of both *MAE* and *RMSE*, Fig. 2 only shows the impacts of γ due to space limitation. From Fig. 2, we can see no matter which training data setting is used, *MAE* and *RMSE* decrease when γ increases. But *MAE* and *RMSE* start to increase when γ is less than a certain threshold such as 0.1. Therefore, setting $\gamma = 0.1$ is proper. So is $\lambda_1 = \lambda_2 = 0.01$.

5.5 Experimental Results

For model validation, we have conducted repeated random sub-sampling for 10 times in each experiment. Finally, each model is experimented with 300 times (3 different percentages $\times$ 10 different initial matrices $\times$ 10 times cross validations). The experimental results, in the best, average and worst initialization cases, are shown in Table 1.

From the results of the three groups of experiments, we can see that in the best initialization cases, our model improves over MATRI by 11.4%–13.6% in term of *MAE* and by 21.8%–24.0% in term of *RMSE*. In the worst initialization cases, the improvements increase to 45.3%–46.4% in term of *MAE* and 39.6%–41.4% in term of *RMSE*. This result means that our model has better robustness. In other words, it not only performs well with the best initialization but also overcomes the worst initialization situations with slightly lower accuracy. In addition, our model improves MFISR by 49.0%–57.9% in term of *MAE* and by 46.5%–53% in term of *RMSE* in all initialization cases.

Summary: The experimental results have demonstrated that our model significantly outperforms the state-of-the-art models in trust prediction accuracy. This is due to the following reasons. *First*, in our model, both personal trust factors and interpersonal trust factors are taken into account comprehensively. *Second*, personal trust factors (i.e., tendencies) are utilized to produce tendency-reduced trust ratings, based on which, the negative effect of trust tendency is reduced. *Third*, different from personal factors, the weighted sum of all interpersonal trust factors becomes part of regularization in matrix factorization. That means propagated trust, trust rating value similarity and rating distribution similarity are all incorporated in trust prediction.

Table 1: Experiment results

Training%	Cases	Metrics	Ours	MATRI	MFISR
80%	Best	*MAE*	**0.1717**	0.1938	0.4006
		RMSE	**0.2284**	0.3004	0.4856
	Average	*MAE*	**0.1802**	0.3091	0.3711
		RMSE	**0.2404**	0.3875	0.4561
	Worst	*MAE*	**0.1883**	0.3514	0.4476
		RMSE	**0.2474**	0.4222	0.5268
60%	Best	*MAE*	**0.1734**	0.1970	0.3578
		RMSE	**0.2362**	0.3022	0.4418
	Average	*MAE*	**0.1804**	0.3109	0.3774
		RMSE	**0.2413**	0.3889	0.4611
	Worst	*MAE*	**0.1862**	0.3476	0.3998
		RMSE	**0.2471**	0.4190	0.4827
40%	Best	*MAE*	**0.1792**	0.2073	0.3516
		RMSE	**0.2389**	0.3099	0.4517
	Average	*MAE*	**0.1821**	0.3165	0.3813
		RMSE	**0.2431**	0.3930	0.4643
	Worst	*MAE*	**0.1855**	0.3392	0.3924
		RMSE	**0.2481**	0.4111	0.4749

6 Conclusion and Future Work

In this paper, we have proposed a trust prediction model based on rating decomposition and matrix factorization. Our model incorporates both personal properties and interpersonal properties in different ways. The personal properties (trust tendencies) are used to decompose trust ratings into truster tendency, trustee tendency and tendency-reduced trust ratings, which reduced the effect of trust tendency. The interpersonal properties (propagated trust and similarities) are incorporated into a propagation and similarity regularization term by which we modified the matrix factorization method to predict trust ratings from tendency-reduced ratings. The experimental results show that this new model outperforms the state-of-the-art trust prediction models by up to 13.6% in term of *MAE* and 24.0% in term of *RMSE*.

In the future, we plan to study social context-aware trust prediction, where similarity can be extended to social context to further improve trust prediction accuracy.

7 Acknowledgments

This work is partially supported by NSFC grant 61303019, and Doctoral Fund of Ministry of Education of China 20133201120012.

References

Berscheid, E., and Reis, H. T. 1998. Attraction and close relationships. *The handbook of social psychology* 2:193281.

Breese, J. S.; Heckerman, D.; and Kadie, C. 1998. Empirical analysis of predictive algorithms for collaborative filtering. In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, 43–52.

Golbeck, J., and Hendler, J. A. 2006. Inferring binary trust relationships in web-based social networks. *ACM Transactions on Internet Technology* 6(4):497–529.

Guha, R.; Kumar, R.; Raghavan, P.; and Tomkins, A. 2004. Propagation of trust and distrust. In *Proceedings of the 13th International Conference on World Wide Web*, 403–412.

Hang, C.-W.; Wang, Y.; and Singh, M. P. 2009. Operators for propagating trust and their evaluation in social networks. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems*, 1025–1032.

Jia, D.; Zhang, F.; and Liu, S. 2013. A robust collaborative filtering recommendation algorithm based on multidimensional trust model. *JSW* 8(1):11–18.

Jøsang, A., and Ismail, R. 2002. The beta reputation system. In *Proceedings of the 15th Bled Conference on Electronic Commerce*.

Koller, D., and Friedman, N. 2009. *Probabilistic Graphical Models: Principles and Techniques*. MIT Press.

Koren, Y.; Bell, R.; and Volinsky, C. 2009. Matrix factorization techniques for recommender systems. *Computer* 42(8):30–37.

Li, L.; Wang, Y.; and Lim, E.-P. 2009. Trust-oriented composite service selection and discovery. In Baresi, L.; Chi, C.-H.; and Suzuki, J., eds., *ICSOC/ServiceWave*, volume 5900 of *Lecture Notes in Computer Science*, 50–67.

Lin, J.; Li, Z.; Wang, D.; Salamatian, K.; and Xie, G. 2012. Analysis and comparison of interaction patterns in online social network and social media. In *21st International Conference on Computer Communications and Networks*, 1 –7.

Liu, G.; Wang, Y.; Orgun, M. A.; and Lim, E.-P. 2013. Finding the optimal social trust path for the selection of trustworthy service providers in complex social networks. *IEEE Transactions on Services Computing* 6(2):152–167.

Liu, G.; Wang, Y.; and Orgun, M. 2009. Trust inference in complex trust-oriented social networks. In *International Conference on Computational Science and Engineering, 2009. CSE '09.*, volume 4, 996–1001.

Liu, G.; Wang, Y.; and Orgun, M. A. 2010. Optimal social trust path selection in complex social networks. In *AAAI*, 1391–1398.

Liu, G.; Wang, Y.; and Orgun, M. A. 2012. Social context-aware trust network discovery in complex contextual social networks. In *AAAI*, 101–107.

Ma, H.; Zhou, D.; Liu, C.; Lyu, M. R.; and King, I. 2011a. Recommender systems with social regularization. In *Proceedings of the fourth ACM international conference on Web search and data mining*, WSDM '11, 287–296. ACM.

Ma, H.; Zhou, T. C.; Lyu, M. R.; and King, I. 2011b. Improving recommender systems by incorporating social contextual information. *ACM Transactions on Information Systems* 29(2):9:1–9:23.

Mitchell, M. 1998. *An Introduction to Genetic Algorithms*. Cambridge, MA, USA: MIT Press.

Salakhutdinov, R., and Mnih, A. 2008a. Bayesian probabilistic matrix factorization using markov chain monte carlo. In *Proceedings of the 25th International Conference on Machine Learning*, 880–887.

Salakhutdinov, R., and Mnih, A. 2008b. Probabilistic matrix factorization. In *Advances in Neural Information Processing Systems*, volume 20.

Tversky, A., and Kahneman, D. 1974. Judgment under uncertainty: Heuristics and biases. *Science* 185:1124–31.

Wang, Y., and Lin, K. J. 2008. Reputation-oriented trustworthy computing in e-commerce environments. *Internet Computing, IEEE* 12(4):55–59.

Wang, Y.; Li, L.; and Liu, G. 2013. Social context-aware trust inference for trust enhancement in social network based recommendations on service providers. *World Wide Web* 1–26.

Xiong, L., and Liu, L. 2004. Peertrust: Supporting reputation-based trust for peer-to-peer electronic communities. *IEEE Transactions on Knowledge and Data Engineering* 16(15):843–857.

Yao, Y.; Tong, H.; Yan, X.; Xu, F.; and Lu, J. 2012. Matrust: An effective multi-aspect trust inference model. *CoRR* abs/1211.2041.

Yao, Y.; Tong, H.; Yan, X.; Xu, F.; and Lu, J. 2013a. Matri: a multi-aspect and transitive trust inference model. In *Proceedings of the 22nd international conference on World Wide Web*, 1467–1476.

Yao, Y.; Tong, H.; Yan, X.; Xu, F.; and Lu, J. 2013b. Multi-aspect + transitivity + bias: An integral trust inference model. *IEEE Transactions on Knowledge and Data Engineering*. PrePrint.

Zhang, H.; Wang, Y.; and Zhang, X. 2012a. Efficient contextual transaction trust computation in e-commerce environments. In *IEEE International Conference on Trust, Security and Privacy in Computing and Communications*, 318–325.

Zhang, H.; Wang, Y.; and Zhang, X. 2012b. A trust vector approach to transaction context-aware trust evaluation in e-commerce and e-service environments. In *SOCA*, 1–8.

Zhang, H.; Wang, Y.; and Zhang, X. 2013. The approaches to contextual transaction trust computation in e-commerce environments. *Security and Communication Networks* Preprint.

Zheng, X.; Wang, Y.; and Orgun, M. 2013. Modeling the dynamic trust of online service providers using HMM. In *IEEE 20th International Conference on Web Services*, 459–466.